Pergamon

Insect Biochemistry and Molecular Biology 30 (2000) 1027–1035

Insect Biochemistry and Molecular Biology

www.elsevier.com/locate/ibmb

Trypsinogen-like cDNAs and quantitative analysis of mRNA levels from the Indianmeal moth, *Plodia interpunctella*

Yu-Cheng Zhu, Karl J. Kramer, Alan K. Dowdy, James E. Baker *

Grain Marketing and Production Research Center, ARS-USDA, 1515 College Avenue, Manhattan, Kansas 66502, USA

Received 2 December 1999; received in revised form 15 March 2000; accepted 16 March 2000

Abstract

Two cDNA fragments encoding full-length trypsinogen-like proteins were cloned from larvae of two strains (RC688s and HD198r) of the Indianmeal moth, *Plodia interpunctella* (Hübner), which differed in their sensitivity to *Bacillus thuringiensis* protoxins. One cDNA fragment contained 874 nucleotides, including a 780-nucleotide open reading frame that encoded a trypsinogen-like protein (*Pi*T2b). Another cDNA fragment amplified from both *P. interpunctella* strains contained 864 nucleotides including a 780 bp open reading frame encoding a second trypsinogen-like protein (*Pi*T2c). The cDNA sequence of *Pi*T2b shared 89% sequence identity with *Pi*T2a, a trypsinogen-like protein cloned previously from this species. The cDNA sequences of *Pi*T2a and *Pi*T2c shared 83% identity. The cDNA sequence identity between *Pi*T2b and *Pi*T2c was 80%. The cDNA for *Pi*T2b from strain RC688s was different at six nucleotide positions from that of *Pi*T2b from strain HD198r. Five nucleotide replacements occurred in the open reading frame leading to amino acid changes at all five positions. There were five nucleotide differences in the cDNAs for *Pi*T2c trypsinogen-like proteins from the two strains. Two nucleotide substitutions in the open reading frame resulted in replacements of two amino acid residues in the deduced protein sequences. Amino acid sequences for *Pi*T2a and *Pi*T2b shared 84% identity, but only 50% identity was observed between *Pi*T2c and the other two trypsinogen-like proteins. The deduced amino acid sequences for *Pi*T2b and *Pi*T2c included both signal and zymogen activation peptides and amino acid sequence motifs which are conserved in seven homologous trypsinogen-like proteins from other insects. Typical features of the putative trypsinogen-like proteins from *P. interpunctella* included the serine proteinase active site triad (His^{81}, Asp^{133}, and Ser^{233}), three pairs of cysteine residues for disulfide bridges, and three residues, Asp^{227}, Gly^{250}, and Gly^{260}, that help to confer trypsin-like specificity to the enzymes. Quantitative RT-PCR analyses showed that, in fourth instar larvae, RC688s had 1.6-fold higher *Pi*T2a trypsinogen-like mRNA than did HD198r. Expression of *Pi*T2b mRNA was 3.4-fold higher in HD198r than in RC688s. Expression of *Pi*T2c mRNA was 2.8-fold higher in RC688s than in HD198r. Mean accumulation levels of mRNAs for all three trypsinogen-like proteins were slightly higher in RC688s than in HD198r based on total RNA, and 1.3-fold higher in RC688s than in HD198r based on wet weight of larval body tissues. Published by Elsevier Science Ltd.

Keywords: Indianmeal moth; *Plodia interpunctella*; *Bacillus thuringiensis*; Midgut; Trypsinogen; Trypsin; Serine proteinases; cDNA; Gene; mRNA; Amino acid sequence; Expression; Stored-product pest; Lepidoptera

1. Introduction

Serine proteinases are common luminal enzymes in midguts of many insect species including Lepidoptera (Applebaum, 1985; Terra and Ferriera, 1994). These endoproteinases function by hydrolysing ingested protein into peptides during the initial stages of protein catabolism. Because of their important biochemical roles in insect growth and development, these enzymes have been targets for proteinaceous inhibitors whose genes have been incorporated into transformed plants (Hilder and Boulter, 1999). Serine proteinases in the midgut also activate protoxins elaborated by the entomopathogen, *Bacillus thuringiensis* (Bt), thereby mediating Bt toxicity (Milne and Kaplan, 1993; Martínez-Ramírez and Real, 1996; Oppert et al., 1996), and they also may play a concurrent role in the hydrolytic degradation and subsequent inactivation of the toxic protein. Forcada et al. (1996) reported that midgut extracts from a Bt-resistant

* Corresponding author. Tel.: +1-785-776-2785; fax: +1-785-537-5584.

E-mail address: baker@usgmrl.ksu.edu (J.E. Baker).

0965-1748/00/$ - see front matter. Published by Elsevier Science Ltd.
PII: S0965-1748(00)00073-4

strain of *Heliothis virescens* degraded the Bt toxin in vitro at a rate faster than extracts from a susceptible strain. Shao et al. (1998) proposed that the excessive degradation of protoxin in *H. armigera* midgut juice, which contained trypsin-, chymotrypsin-, and elastase-like enzymes, was responsible for the low sensitivity of this species to Bt, with chymotrypsin playing a major role in the degrading process.

Plants transformed with genes for Bt toxins are targeted to disrupt gut epithelial cell physiology but not to inhibit proteinases. However, the potential for gut proteinases to modulate the effectiveness of Bt protoxins in insects is high and was the impetus for more detailed studies on the properties and gene expression of these enzymes in the Indianmeal moth, *Plodia interpunctella* (Hübner), a pest of stored products that has developed resistance to Bt (McGaughey, 1985). There is evidence that, in a selected strain of *P. interpunctella* (HD198r), reduced sensitivity to Bt is associated with the absence of a proteolytic activity that is present in the wild-type strain (RC688s) (Oppert et al., 1997). The pattern of proteolytic cleavage of Bt protoxins differs significantly between these strains (Oppert et al. 1994, 1996). However, despite the differences in midgut proteinase activities, the initial cloning and characterization of the genes for these enzymes at the molecular level documented rather similar cDNA sequences and mRNA expression levels for both chymotrypsinogen-like and trypsinogen-like genes in the two strains (Zhu et al. 1997, 2000). Nevertheless, in the latter study, Southern blot analysis revealed differences in restriction enzyme cleavage sites in the trypsinogen-like genes from the two *Plodia* strains, which in addition to other data, suggested the possibility of the presence of multiple trypsinogen-like genes in this species. The purpose of the current study was to determine if additional trypsinogens are present in these two *Plodia* strains and to apply a quantitative PCR method for monitoring the expression levels of their corresponding mRNAs.

2. Materials and methods

2.1. Insect cultures

P. interpunctella strain RC688s was collected from farm-stored grain in Riley County, Kansas, and maintained on a diet described by McGaughey and Beeman (1988). A Bt-resistant strain, HD198r, was selected from RC688s using Bt subsp. *entomocidus*, strain HD-198 (McGaughey and Johnson, 1992). The resistance level of strain HD198 to Bt subsp. *entomocidus* is 100-fold higher than that of strain RC688 (McGaughey, 1985).

2.2. Cloning trypsinogen-like protein cDNAs

Messenger RNA was purified from both strains of *P. interpunctella* larvae. cDNA was synthesized and directionally cloned into Uni-ZAP phage vector (Stratagene). In the previous study, one trypsin-like cDNA was isolated and sequenced from both *P. interpunctella* strains (Zhu et al., 2000) by using traditional library screening protocols. To obtain more highly-sequence-specific trypsinogen-like cDNAs, polymerase chain reaction (PCR) was conducted to amplify trypsinogen-like cDNAs using the cDNA library as template and a T7 reverse vector primer and a forward degenerate primer, 5′-TGY-CARGGNGAYWSNGGNGGNCCNYT-3′, designed from a highly conserved region (CQGDSGGPL) in both *Manduca sexta* trypsin and chymotrypsin that is located approximately 250 bp from the 3′-end of these genes (Peterson et al. 1994, 1995). PCR-amplified DNA fragments (≈250 bp) were cloned into a pGEM-T vector (Promega, Madison, WI). Sequences of these clones were determined by using an automated sequencer. The cDNA sequence of the trypsinogen-like protein was confirmed by homology searching of GenBank provided by the National Center for Biotechnology Information by using the BLASTX protocol (Altschul et al., 1990; Gish and States, 1993).

After the 3′-end sequence of the trypsinogen-like cDNA was determined, specific reverse primers for each trypsin-like cDNA were designed. The 5′ ends of different trypsin-like cDNAs were amplified using a T3 vector forward primer and a reverse primer designed from the sequence of each trypsin cDNA fragment. The PCR fragments were cloned into the pGEM-T vector and the sequences determined.

To specifically amplify the full sequence of each trypsinogen-like cDNA, a forward primer and a reverse primer were designed based on sequence information of each trypsinogen-like cDNA. PCR amplification was performed using a thermostable proofreading Pfu DNA polymerase (Promega) to re-amplify a full length trypsinogen-like cDNA fragment from the cDNA library template. This fragment was A-tailed and cloned into the pGEM-T vector. The sequence of the insert was determined from both directions.

The Wisconsin Sequence Analysis Package (GCG Unix version 9.0, Genetics Computer Group, Madison, WI) including Pileup, Gap, Distances, and Growtree programs was used to analyse the similarity of trypsinogen-like protein sequences (gap weight=2, gap length weight=1). Sequence analysis tools of the ExPASy Molecular Biology Server of Swiss Institute of Bioinformatics were used to process data of deduced protein sequences.

2.3. Quantitative reverse transcription (QRT) PCR analysis of trypsinogen-like mRNA expression

Procedures for the quantitative determination of the expression levels of three trypsinogen-like mRNAs were modified from those used by Alexandre et al. (1998), Guenthner and Hart (1998), and Igaz et al. (1998). To develop a homogeneous internal standard for QRT-PCR analysis of trypsinogen *Pi*T2a mRNA expression, the restriction enzyme *Bam* HI was used to cut out a fragment of *Pi*T2a cDNA (nucleotide 268–291, Zhu et al., 2000). The fragments were religated using T4 DNA ligase. Because of two identical sticky ends cut by *Bam* HI, re-ligation resulted in a cDNA fragment with a deletion or insertion of different numbers of 24-bp fragments. A cDNA fragment with two 24-bp insertions was identified by PCR and agarose gel electrophoresis, and then subsequently used as the internal standard for QRT-PCR analysis of *Pi*T2a mRNA expression. Similarly, restriction enzyme *Hinf* I (Fig. 1) was used to produce an internal standard with a 138-bp insertion (containing two 69-bp fragments) for QRT-PCR analysis of trypsinogen *Pi*T2b mRNA expression. Restriction enzyme *Aat* II (Fig. 2) was used to produce an internal standard with a 125-bp deletion for QRT-PCR analysis of trypsinogen *Pi*T2c mRNA expression.

Total RNA from a group of 20 fourth instar larvae (~70 mg) fed on the standard laboratory diet was extracted by using guanidine thiocyanate denaturing solution and precipitated with isopropanol (Titus, 1991). Total RNA extractions were repeated three times for both Bt-susceptible and Bt-resistant strains of *P. interpunctella*. Reverse transcription was conducted with an oligo-dT primer and SuperScript II reverse transcriptase (Gibco BRL Life-Technologies, Gaithersburg, MD). Fifteen micrograms of RNA for each repeat were used to generate the first strand of cDNA.

QRT-PCR was conducted three times with different RNA samples for each trypsinogen-like mRNA species and insect strain. Specific primers were designed based on unique sequences for each trypsinogen-like cDNA. Primers T2F and T2R (sequences are underlined in Fig. 1) were used to amplify *Pi*T2b fragment, and primers T3F and T3R (Fig. 2) were used to amplify *Pi*T2c fragment. The concentration range of the internal standard was determined by preliminary experiments. The standard was diluted in distilled H_2O as follows: 10,000; 5,000; 1,000; 500; 100; 50; 10; 5; and 1 pg/μl. The PCR reaction mixture contained 1.5 μl of standard, 1.5 μl of target cDNA, 2 μl of PCR buffer (100 mM Tris–chloride pH 9.0 containing 500 mM KCl and 15 mM $MgCl_2$), 0.8 μl of 2.5 mM of each dNTP in H_2O, 0.5 μl of each primer (10 pmole/μl H_2O), 1.5 unit of Taq DNA polymerase, and H_2O to obtain a final volume of 20 μl. The DNA was initially denatured for 3 min at 94°C, and the PCR amplification included 40 cycles of 30 s denaturing at 94°C, 30 s annealing at 57°C, and 1 min extension at 72°C in a PTC-100 thermocycler (MJ Research, Inc., Watertown, MA). PCR products were separated on 1% agarose gels, stained with 0.5 μg/ml ethidium bromide, and photographed under UV light. Photographic data was analysed by using the video densitometry program ITTI version 2.2 (Interactive Technologies International). The relationship between standard concentration and PCR band intensity was determined by using Sigma Plot with a log transform of the concentration.

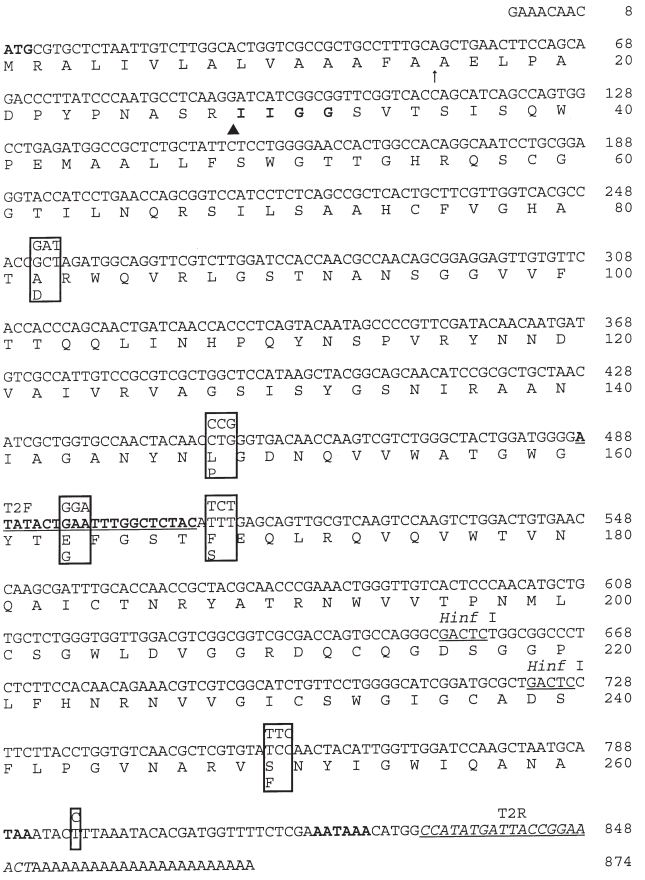

Fig. 1. Nucleotide and deduced amino acid sequences of trypsinogen-like cDNA *Pi*T2b prepared from *P. interpunctella* RNA (strain RC688s). ATG=start codon; TAA=termination codon; AATAAA=polyadenylation signal; ↑=predicted signal peptide cleavage site; ▲=predicted activation peptide cleavage site. IIGG are potential N-terminal residues. Different nucleotides (on the top) and predicted amino acid residues (at the bottom) of HD198r strain are boxed at each corresponding position. Sequences of 2 *Hinf* I sites were underlined. T2F primer sequence is marked with underlined bold letters, and T2R primer sequence is marked with underlined italic letters. The sequence has been deposited in GenBank with accession numbers AF173495 for RC688s strains and AF173496 for HD198r strain.

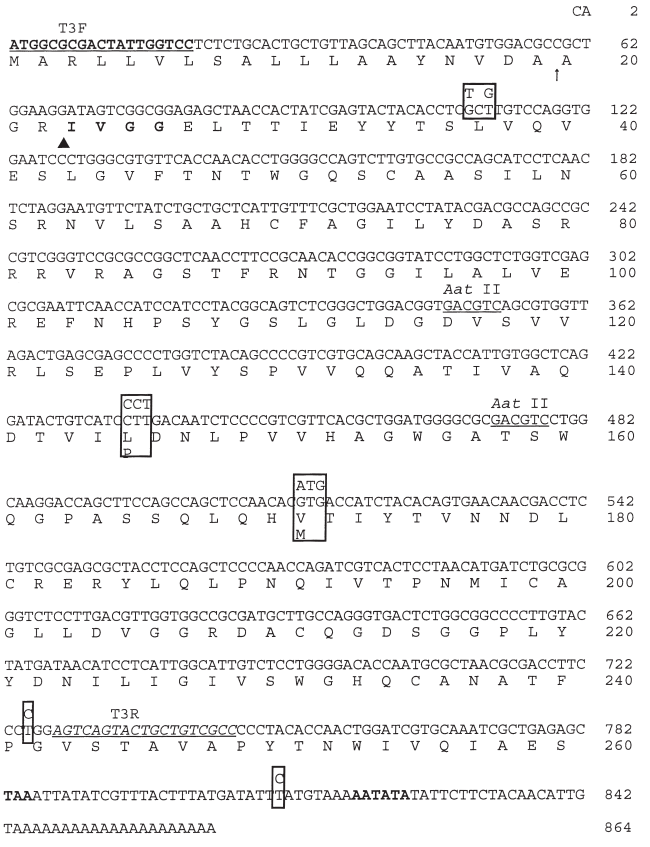

Fig. 2. Nucleotide and deduced amino acid sequences of trypsinogen-like cDNA *Pi*T2c prepared from *P. interpunctella* RNA (strain RC688s). ATG=start codon; TAA=termination codon; AATATA=potential polyadenylation signal; ↑=predicted signal peptide cleavage site; ▲=predicted activation peptide cleavage site. IVGG are conserved N-terminal residues. Different nucleotides (on the top) and predicted amino acid residues (at the bottom) of HD198r strain are boxed at each corresponding position. Sequences of two *Aat* II sites are underlined. T3F primer sequence is marked with underlined bold letters, and T3R primer sequence is marked with underlined italic letters. The sequences have been deposited in GenBank with accession number AF173497 for RC688s strain and AF173498 for HD198r strain.

3. Results

3.1. Trypsinogen-like protein cDNAs

3′-RACE was used to obtain cDNA fragments encoding putative trypsinogens from *P. interpunctella*. Sixteen of 31 clones that were sequenced carried inserts with deduced amino acid sequences similar to other trypsinogen-like proteins in the GenBank, whereas only two clones carried inserts with a deduced amino acid sequence similar to chymotrypsinogen-like proteins. Multiple alignment of these trypsinogen-like cDNA ends demonstrated that these sequences represented three different cDNAs including the one reported previously, hereafter termed *Pi*T2a (Zhu et al., 2000). The 5′-end of each trypsinogen-like cDNA was amplified and sequenced using specific primers designed for each kind of trypsinogen based on the previously obtained 3′-cDNA end sequences. Full-length cDNA fragments were successfully amplified from the cDNA libraries of both strains by using Pfu DNA polymerase and specific primers corresponding to 5′- and 3′-ends of each trypsinogen-like cDNA.

In addition to the cDNA coding for *Pi*T2a, two additional full-length sequence cDNA fragments were cloned from both *P. interpunctella* strains. One cDNA fragment contained 874 nucleotides that encoded a trypsinogen-like protein (designated as *Pi*T2b). The cDNA sequences from both strains included the start codon ATG at positions 9–11, the termination codon TAA at positions 789–791, and the polyadenylation signal, AATAAA, at positions 821–826 (Fig. 1). The open reading frame consisted of 780 nucleotides. Another cDNA fragment amplified from both strains encoded a third trypsinogen-like protein (designated as *Pi*T2c). The sequences included a start codon, a termination codon, a potential polyadenylation signal sequence AATATA, and a 780-bp open reading frame (Fig. 2).

The cDNA sequences for *Pi*T2a and *Pi*T2b, *Pi*T2a and *Pi*T2c, and *Pi*T2b and *Pi*T2c shared 89%, 83%, and 80% identities, respectively.

3.2. Deduced protein sequences

The predicted amino acid sequence from the *Pi*T2b cDNA contained 260 residues. The deduced mature trypsin-like protein had 232 amino acid residues including nine negatively-charged residues (aspartic acid and glutamic acid) and 13 positively charged residues (arginine and lysine). The molecular mass and theoretical pI value were 25 kDa and pH 8.78, respectively. The deduced trypsinogen-like protein from the *Pi*T2c cDNA contained 260 amino acid residues. The mature form of *Pi*T2c contained 238 amino acid residues with 18 negatively-charged residues and 11 positively-charged residues. The molecular mass and theoretical pI value were 26 kDa and pH 5.05, respectively.

The deduced amino acid sequence of *Pi*T2b was most similar to *Pi*T2a with 84% identity (Fig. 3). Less similar

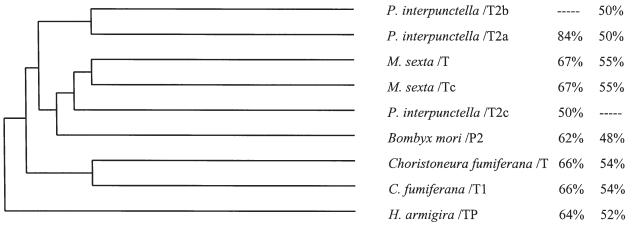

Fig. 3. Phylogenetic relationship based on the amino acid sequences of *Pi*T2a, *Pi*T2b, and *Pi*T2c from *P. interpunctella* (RC688s) and 6 other homologous trypsinogen sequences (see GenBank citations in Fig. 4) established using GCG Distances and Growtree programs. Sequence identity with *Pi*T2b is displayed in the first column after species name, and sequence identity with *Pi*T2c is displayed in the second column after species name.

were the sequences of the trypsin-like precursors T and Tc from *Manduca sexta* (both 67%). The sequence identity between *Pi*T2b and *Pi*T2c was 50%. Gap analysis indicated that the *Pi*T2c sequence was most similar to those of two *M. sexta* trypsinogens with 55% sequence identity.

Multiple sequence alignment of the three trypsinogen-like proteins of *P. interpunctella* with six homologous insect trypsin-like proteases indicated that 71 amino acid residues were conserved among all nine of the insect trypsin-like proteins (Fig. 4). These sequences contained all of the conserved residues representing typical features for trypsin-like proteinases, including signal and zymogen activation peptides, a highly conserved N-terminal sequence, three conserved putative active site residues, His^{81}, Asp^{133}, and Ser^{233}, which form the catalytic triad in serine proteases (Kraut, 1977; Wang et al., 1993; Peterson et al., 1994), and six cysteine residues predicted to occur in disulfide bridge configurations among trypsins and chymotrypsins. The residues Asp^{227}, Gly^{250}, and Gly^{260}, which help to define the substrate binding pocket are highly conserved in these trypsin-like enzymes as well. Asp^{227} determines specificity in both invertebrate and vertebrate trypsins by interacting through ionic forces with a Lys or Arg residue at the substrate cleavage site (Hedstrom et al., 1992; Wang et al., 1993).

3.3. Differences in cDNAs and deduced protein sequences

The cDNA for *Pi*T2b of strain RC688s was different at six nucleotide positions from that of strain HD198r. Five nucleotide replacements were present on the open reading frame (Fig. 1), leading to amino acid changes at all five of these positions. Nucleotide substitutions, $C^{245}\rightarrow A^{245}$, $T^{423}\rightarrow C^{423}$, $A^{488}\rightarrow G^{488}$, $T^{503}\rightarrow C^{503}$, $C^{749}\rightarrow T^{749}$, in strain RC688s resulted in amino acid changes, $A^{82}\rightarrow D^{82}$, $L^{148}\rightarrow P^{148}$, $E^{163}\rightarrow G^{163}$, $F^{168}\rightarrow S^{168}$, $S^{250}\rightarrow F^{250}$, in strain HD198r. Despite the differences in amino acid content, the mature trypsin-like enzyme of HD198r strain was still predicted to have a molecular mass of 25 kDa, a pI value of pH 8.78, and the same numbers of negatively charged residues (9) and positively charged residues (13) as was predicted from RC688s cDNA.

There were five nucleotide differences on cDNAs from RC688s and HD198r strains for *Pi*T2c trypsinogen-like proteins, including two nucleotide substitutions in the open reading frame, which resulted in replacements of two amino acid residues in the deduced amino acid sequences (Fig. 2). None of the different amino acid residues resulted in any differences in molecular mass, pI value, or numbers of negatively- or positively-charged residues between the two strains.

Among three trypsinogen-like proteins from *Plodia*, *Pi*T2a and *Pi*T2b were very similar, with 89% cDNA sequence identity and 84% amino acid sequence identity. *Pi*T2c is less similar and more closely related to two *Manduca* trypsin-like enzyme precursors than to the other two trypsinogen-like proteins from *Plodia* (Figs 3 and 4).

3.4. Trypsinogen mRNA expression

Different levels of trypsinogen-like mRNA expression in the two *P. interpunctella* strains were first observed after PCR amplification of equal titers of RC688s and HD198r cDNA libraries by using specific primers for each trypsinogen-like cDNA. Band intensities for *Pi*T2a and *Pi*T2c amplified from RC688s were similar to those from strain HD198r (indicated by arrows in Fig. 5a). DNA fragment intensity for *Pi*T2b amplified from the RC688s library was 3.1-fold lower than that for *Pi*T2b from HD198r. The results were reproducible when different primers were used in the PCR amplifications (data not shown).

By using quantitative RT-PCR, expression levels of each trypsinogen-like mRNA in fourth instar larvae of *P. interpunctella* were determined. Densitometric analysis showed that the total area of each band reached equilibrium with the internal standard at concentrations of 7.5 pg mRNA per μl of PCR product for RC688s and 4.4 pg mRNA per μl of PCR product for HD198r. Strain RC688s had 1.2-fold higher *Pi*T2a trypsinogen-like mRNA than strain HD198r based on mRNA content per μg of total RNA and 1.6-fold higher *Pi*T2a mRNA than HD198r based on mRNA content per gram of larval tissue (*Pi*T2a, Fig. 5b). Quantitative analyses of mRNA for *Pi*T2b proteins in the two strains indicated that strain HD198r had 4.7-fold higher *Pi*T2b mRNA per μg of total RNA and 3.4-fold higher *Pi*T2b mRNA per gram tissue (*Pi*T2b, Fig. 5b) than strain RC688s. Results of QRT-PCR analyses on *Pi*T2c mRNA indicated that *Pi*T2c mRNA expression levels in strain RC688s were 2-fold higher based on total RNA and 2.8-fold higher based on tissue weight (*Pi*T2c, Fig. 5b) than strain HD198r.

Although the three trypsinogen-like mRNAs were expressed differently in the two *P. interpunctella* strains, overall accumulative levels of mRNAs for the three trypsinogen-like proteins in larvae fed the laboratory diet were only slightly higher (1.1-fold) in RC688s than in HD198r based on mRNA levels per μg of total RNA, and 1.3-fold higher based larval tissue weight (PiT2a+b+c; Fig. 5b).

4. Discussion

The proteinase from insects studied most frequently has been trypsin-like in specificity (Reeck et al., 1999). We have studied trypsinogen-like proteins in *P. inter-*

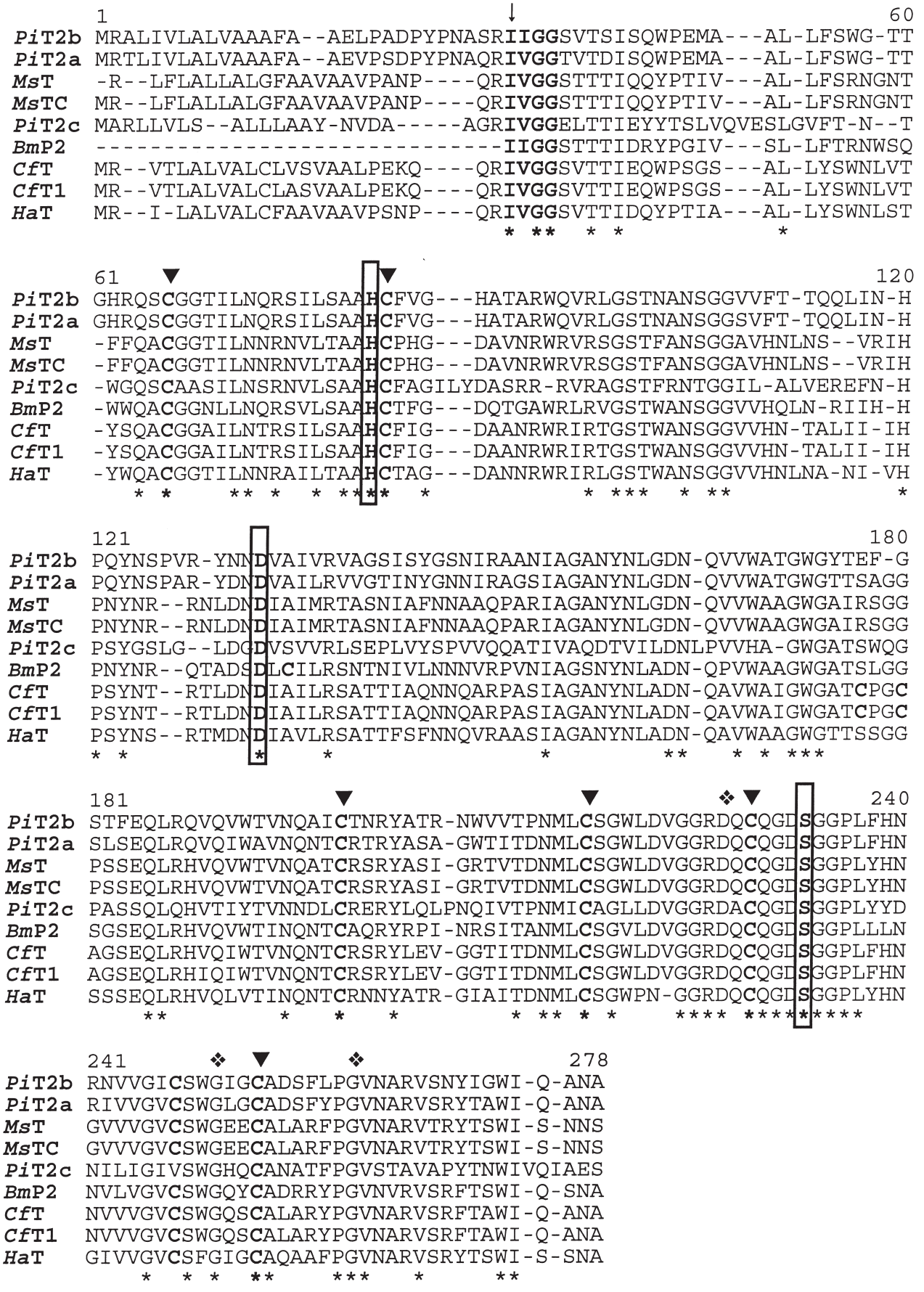
1 60
PiT2b MRALIVLALVAAAFA--AELPADPYPNASRIIGGSVTSISQWPEMA---AL-LFSWG-TT
PiT2a MRTLIVLALVAAAFA--AEVPSDPYPNAQRIVGGTVTDISQWPEMA---AL-LFSWG-TT
MsT -R--LFLALLALGFAAVAAVPANP----QRIVGGSTTTIQQYPTIV---AL-LFSRNGNT
MsTC MR--LFLALLALGFAAVAAVPANP----QRIVGGSTTTIQQYPTIV---AL-LFSRNGNT
PiT2c MARLLVLS--ALLLAAY-NVDA-----AGRIVGGELTTIEYYTSLVQVESLGVFT-N--T
BmP2 ----------IIGGSTTTIDRYPGIV---SL-LFTRNWSQ
CfT MR--VTLALVALCLVSVAALPEKQ----QRIVGGSVTTIEQWPSGS---AL-LYSWNLVT
CfT1 MR--VTLALVALCLASVAALPEKQ----QRIVGGSVTTIEQWPSGS---AL-LYSWNLVT
HaT MR--I-LALVALCFAAVAAVPSNP----QRIVGGSVTTIDQYPTIA---AL-LYSWNLST
61 120
PiT2b GHRQSCGGTILNQRSILSAAHCFVG---HATARWQVRLGSTNANSGGVVFT-TQQLIN-H
PiT2a GHRQSCGGTILNQRSILSAAHCFVG---HATARWQVRLGSTNANSGGSVFT-TQQLIN-H
MsT -FFQACGGTILNNRNVLTAAHCPHG---DAVNRWRVRSGSTFANSGGAVHNLNS--VRIH
MsTC -FFQACGGTILNNRNVLTAAHCPHG---DAVNRWRVRSGSTFANSGGAVHNLNS--VRIH
PiT2c -WGQSCAASILNSRNVLSAAHCFAGILYDASRR-RVRAGSTFRNTGGIL-ALVEREFN-H
BmP2 -WWQACGGNLLNQRSVLSAAHCTFG---DQTGAWRLRVGSTWANSGGVVHQLN-RIIH-H
CfT -YSQACGGAILNTRSILSAAHCFIG---DAANRWRIRTGSTWANSGGVVHN-TALII-IH
CfT1 -YSQACGGAILNTRSILSAAHCFIG---DAANRWRIRTGSTWANSGGVVHN-TALII-IH
HaT -YWQACGGTILNNRAILTAAHCTAG---DANNRWRIRLGSTWANSGGVVHNLNA-NI-VH
121 180
PiT2b PQYNSPVR-YNNDVAIVRVAGSISYGSNIRAANIAGANYNLGDN-QVVWATGWGYTEF-G
PiT2a PQYNSPAR-YDNDVAILRVVGTINYGNNIRAGSIAGANYNLGDN-QVVWATGWGTTSAGG
MsT PNYNR--RNLDNDIAIMRTASNIAFNNAAQPARIAGANYNLGDN-QVVWAAGWGAIRSGG
MsTC PNYNR--RNLDNDIAIMRTASNIAFNNAAQPARIAGANYNLGDN-QVVWAAGWGAIRSGG
PiT2c PSYGSLG--LDGDVSVVRLSEPLVYSPVVQQATIVAQDTVILDNLPVVHA-GWGATSWQG
BmP2 PNYNR--QTADSDLCILRSNTNIVLNNNVRPVNIAGSNYNLADN-QPVWAAGWGATSLGG
CfT PSYNT--RTLDNDIAILRSATTIAQNNQARPASIAGANYNLADN-QAVWAIGWGATCPGC
CfT1 PSYNT--RTLDNDIAILRSATTIAQNNQARPASIAGANYNLADN-QAVWAIGWGATCPGC
HaT PSYNS--RTMDNDIAVLRSATTFSFNNQVRAASIAGANYNLADN-QAVWAAGWGTTSSGG
181 240
PiT2b STFEQLRQVQVWTVNQAICTNRYATR-NWVVTPNMLCSGWLDVGGRDQCQGDSGGPLFHN
PiT2a SLSEQLRQVQIWAVNQNTCRTRYASA-GWTITDNMLCSGWLDVGGRDQCQGDSGGPLFHN
MsT PSSEQLRHVQVWTVNQATCRSRYASI-GRTVTDNMLCSGWLDVGGRDQCQGDSGGPLYHN
MsTC PSSEQLRHVQVWTVNQATCRSRYASI-GRTVTDNMLCSGWLDVGGRDQCQGDSGGPLYHN
PiT2c PASSQLQHVTIYTVNNDLCRERYLQLPNQIVTPNMICAGLLDVGGRDACQGDSGGPLYYD
BmP2 SGSEQLRHVQVWTINQNTCAQRYRPI-NRSITANMLCSGVLDVGGRDQCQGDSGGPLLLN
CfT AGSEQLRHVQIWTVNQNTCRSRYLEV-GGTITDNMLCSGWLDVGGRDQCQGDSGGPLFHN
CfT1 AGSEQLRHIQIWTVNQNTCRSRYLEV-GGTITDNMLCSGWLDVGGRDQCQGDSGGPLFHN
HaT SSSEQLRHVQLVTINQNTCRNNYATR-GIAITDNMLCSGWPN-GGRDQCQGDSGGPLYHN
241 278
PiT2b RNVVGICSWGIGCADSFLPGVNARVSNYIGWI-Q-ANA
PiT2a RIVVGVCSWGLGCADSFYPGVNARVSRYTAWI-Q-ANA
MsT GVVVGVCSWGEECALARFPGVNARVTRYTSWI-S-NNS
MsTC GVVVGVCSWGEECALARFPGVNARVTRYTSWI-S-NNS
PiT2c NILIGIVSWGHQCANATFPGVSTAVAPYTNWIVQIAES
BmP2 NVLVGVCSWGQYCADRRYPGVNVRVSRFTSWI-Q-SNA
CfT NVVVGVCSWGQSCALARYPGVNARVSRFTAWI-Q-ANA
CfT1 NVVVGVCSWGQSCALARYPGVNARVSRFTAWI-Q-ANA
HaT GIVVGVCSFGIGCAQAAFPGVNARVSRYTSWI-S-SNA

Fig. 4. Predicted amino acid sequences of three trypsinogen-like proteins from *P. interpunctella* (strain RC688s) and alignment with 6 other insect trypsin-like sequences. *Ms*T=trypsinogen-like sequence (GenBank: AAA39241) of *M. sexta*; *Ms*TC=trypsin C precursor (GenBank: P35047) from *M. sexta*. *Bm*P2=serine proteinase P-II (GenBank: S32398) from the silkworm, *Bombyx mori*. *Cf*T=trypsin (GenBank: AAA84423) from the spruce budworm, *C. fumiferana*. *Cf*T1=trypsin CFT-1 precursor (GenBank: P35042) from the spruce budworm, *C. fumiferana*. *Ha*T=trypsin-like proteinase (GenBank: CAA72950) from the cotton bollworm *H. armigera*. Functionally important residues His^{81}, Asp^{133}, and Ser^{233} are boxed and indicated by bold letters. Cysteines corresponding to the sites of predicted disulfide bridges are marked with bold letters and solid triangle (▼) on the top. Conserved residues for the trypsin binding pocket, Asp^{227}, Gly^{250}, and Gly^{260}, are indicated by (❖) at the top of the sequences. Identical residues among all nine sequences are indicated with stars (★) at the bottom of the sequences. The arrow (↓) indicates the N-terminal residues of the active enzymes. Hyphens represent sequence alignment gaps.

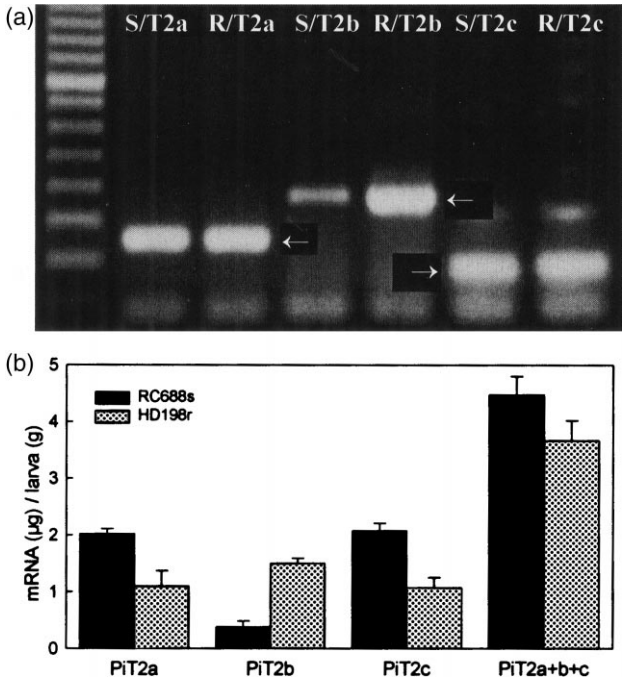

Fig. 5. Quantitative analyses of mRNA expression levels in two *Plodia* strains. (a) PCR amplifications of RC688s (S) and HD198r (R) cDNA libraries by using primers specific for each of three trypsinogen-like cDNAs (*Pi*T2a, b, c) and resultant bands indicated by arrows. (b) mRNA expression levels of all three trypsinogen-like proteins (*Pi*T2a, b, c) in RC688s and HD198r strains. *Pi*T2a+b+c=total amount of mRNA for all three of the trypsinogens.

punctella midguts because this type of proteinase is important in solubilizing and activating the protoxins produced by *B. thuringiensis*, and this species exhibits a trypsin-like proteinase-based mechanism of resistance to Bt (Oppert et al., 1997).

We observed both structural differences in the cDNAs and differences in expression of the individual trypsinogen mRNAs between strains RC688s and HD198r of *P. interpunctella*. The N-terminal sequence of both *Pi*T2a and *Pi*T2c is IVGG, a highly conserved sequence in many trypsin-like proteinases (Wang et al., 1993; Valaitis et al., 1999). *Pi*T2b cDNA had an N-terminus of IIVV which is similar to the N-terminus of a trypsin-like proteinase from the silkworm, *Bombyx mori* (Sasaki et al., 1993). *Pi*T2a and b are very similar in amino acid sequence, whereas *Pi*T2c is less similar. mRNA levels of *Pi*T2a and c were higher in RC688s than in HD198r, whereas *Pi*T2b mRNA was lower. Total trypsinogen mRNA was only 1.1-fold higher in RC688s compared with that of HD198r, which differed from the 2.2-fold higher trypsin-like activity in RC688s compared with that of HD198r (Oppert et al., 1997). However, these latter activity measurements for RC688s also would include the contribution toward hydrolysis of BA*p*NA by the 45 kDa BA*p*NAase in this strain, the mRNA of which was not measured and not included in the total measured mRNA of the cloned trypsinogens in this study. Although our results provide evidence for the production of relatively equal transcriptional levels of 25 kDa trypsin-like proteins in the two strains, the significance of the sequence differences and mRNA level relationships of the individual trypsinogens for the two strains of *Plodia* will require expression and purification of each trypsinogen and detailed studies on their biochemical properties, specificities, and possible dietary effects.

Trypsin-like genes have multiple copies in many insects. Peterson et al. (1994) reported three cDNAs encoding alkaline midgut trypsins in *M. sexta*. Seven trypsinogen cDNAs were identified in *Anopheles gambiae* (Müller et al., 1993). Wang et al. (1995) suggested that ancestors of dipterans and lepidopterans may have had only one trypsinogen gene, but that some species may have gained extra copies by subsequent gene duplications. In *Drosophila*, a relatively complex pattern of molecular evolution for five clustered trypsin genes was proposed, which involved chromosome looping and gene conversion (Wang et al., 1999). Two of these genes apparently evolved in concert, one independently, and the other two genes exhibited an intermediate pattern of evolution. The function of these gene families is unclear. Multigene trypsin families may have evolved to provide a more efficient mechanism for protein digestion as well as to provide an adaptive advantage for phytophagous species feeding on plants that contain proteinase inhibitors (Bown et al., 1997; Reeck et al., 1999). Multigene families may allow the induction of proteinases that are insensitive to dietary inhibitors, as noted by Jongsma et al. (1995), Broadway (1995), and Bolter and Jongsma (1995), or they may initiate proteolysis of proteinase inhibitors by non-target digestive proteinases (Michaud et al., 1995).

In addition to multiple trypsin-like coding regions in *P. interpunctella*, there were also strain differences in the sequences of the enzymes. In *Pi*T2b and *Pi*T2c, there were five and two amino acid replacements, respectively,

in the resistant strain. We compared the deduced sequences of *Pi*T2b and *Pi*T2c with that of human beta-tryptase, a serine proteinase produced in mast cells and thought to cause several allergic disorders (Pereira et al., 1998), by sequence alignment and by examination of three-dimensional structure. In *Pi*T2b, an aliphatic residue Ala^{82} (Fig. 1), corresponding to Ala^{54} of human tryptase located at a position opposite the binding pocket and close to an active site residue histidine, is replaced by an acidic residue Asp^{82} in the resistant strain. In both *Pi*T2b (Fig. 1) and *Pi*T2c (Fig. 2), a leucine in the susceptible strain protein is substituted by proline, corresponding to proline in the human tryptase which is distant from any functional groups. In *Pi*T2b, a glutamic acid at position 163 (Fig. 1) in the susceptible strain is replaced by an aliphatic residue glycine in the resistant strain, corresponding to an aspartic $acid^{134}$ residue in the human tryptase, which is located close to the specificity-determining residue aspartic $acid^{188}$. The remaining two amino acid changes, $F^{168}\rightarrow S^{168}$ and $S^{250}\rightarrow F^{250}$, occurred at some distance from the functional residues (Fig. 1). In *Pi*T2c, a conserved $valine^{171}$ in the susceptible strain is replaced by a methionine, corresponding to residue $valine^{148}$ of human tryptase which is located near the binding pocket (Fig. 2).

The effects of these changes in amino acids on the biochemical function of these trypsin-like enzymes in the *Plodia* strains, including possible altered interaction with Bt protoxins, need to be determined. Numerous attempts to express these trypsins by using a recombinant fusion protein expression system in *Escherichia coli* were unsuccessful because the host cells failed to grow and produce active proteins (Zhu et al., unpublished data).

In Lepidoptera, insects may regulate the midgut proteinase profiles to adapt to differences in protein quality of ingested food or to the presence of proteinase inhibitors in the diet (Broadway and Duffey 1986, 1988). Compared to insects on an inhibitor-free diet, chymotrypsin mRNA in *Helicoverpa armigera* increased in larvae fed diet containing inhibitors, whereas trypsin mRNA levels in these larvae decreased (Gatehouse et al., 1997). Changing the protein level of the diet did not affect trypsin mRNA levels, but chymotrypsin mRNA levels decreased with increasing dietary protein. These non-intuitive results complicated attempts to develop a model of proteinase regulation in this species. In *P. interpunctella*, physiological differences related to fitness costs associated with Bt resistance (Oppert et al., unpublished data) also complicate the characterization of the resistance mechanism(s) at the level of enzyme expression. Nevertheless, we have observed significant differences between the two strains in the regulation of three different trypsinogen-like proteins when these larvae are fed identical diets. Besides structural differences in *Pi*T2b and *Pi*T2c, relative to the susceptible strain, the resistant strain has lower levels of mRNA for *Pi*T2a and *Pi*T2c and higher levels of *Pi*T2b mRNA. The adaptive significance of both the altered protein structures and regulatory processes for trypsinogen-like proteins in *Plodia*, as well as the functional role during the ingestion of Bt toxin or other types of inhibitors, will be determined in the future.

Acknowledgements

The authors thank S. Muthukrishnan and M. Kanost, Kansas State University, and B. Oppert of USDA-ARS-GMPRC, Manhattan KS, for reviewing an earlier version of this manuscript.

References

Alexandre, I., Zammatteo, N., Ernest, I., Ladriere, J.-M., Le, L., Hamels, S., Chandelier, N., Vipond, B., Remacle, J., 1998. Quantitative determination of CMV DNA using a combination of competitive PCR amplification and sandwich hybridization. BioTechniques 25, 676–683.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. J. Molec. Biol. 215, 403–410.

Applebaum, S.W., 1985. Biochemistry of digestion. In: Kerkut, G.A., Gilbert, L.I. (Eds.), Comprehensive Insect Physiology, Biochemistry and Pharmacology, vol. 4. Pergamon Press, New York, pp. 279–311.

Bolter, C.J., Jongsma, M.A., 1995. Colorado potato beetles (*Leptinotasa decemlineata*) adapt to proteinase inhibitors induced in potato leaves by methyl jasmonate. J. Insect Physiol. 41, 1071–1078.

Bown, D.P., Wilkinson, H.S., Gatehouse, J.A., 1997. Differentially regulated inhibitor-sensitive and insensitive protease genes from phytophagous insect pest. *Helicoverpa armigera*, are members of complex multigene families. Insect Biochem. Molec. Biol. 27, 625–638.

Broadway, R.M., Duffey, S.S., 1986. Plant proteinase inhibitors: mechanism of action and effect on the growth and digestive physiology of larval *Heliothis zea* and *Spodoptera exigua*. J. Insect Physiol. 32, 827–834.

Broadway, R.M., Duffey, S.S., 1988. The effect of plant protein quality on insect digestive physiology and the toxicity of plant proteinase inhibitors. J. Insect Physiol. 34, 1111–1117.

Broadway, R.M., 1995. Are insects resistant to plant proteinase inhibitors? J. Insect Physiol. 41, 107–116.

Forcada, C., Alcácer, E., Garcera, M.D., Martinez, R., 1996. Differences in the midgut proteolytic activity of two *Heliothis virescens* strains, one susceptible and one resistant to *Bacillus thuringiensis* toxins. Arch. Insect Biochem. Physiol. 31, 257–272.

Gatehouse, L.N., Shannon, A.L., Burgess, E.P.J., Christeller, J.T., 1997. Characterization of major midgut proteinase cDNAs from *Helicoverpa armigera* larvae and changes in gene expression in response to four proteinase inhibitors in the diet. Insect Biochem. Molec. Biol. 27, 929–944.

Gish, W., States, D.J., 1993. Identification of protein coding regions by database similarity search. Nature Genet. 3, 266–272.

Guenthner, P.C., Hart, C.E., 1998. Quantitative, competitive PCR assay for HIV-1 using a microplate-based system. BioTechniques 24, 810–816.

Hedstrom, L., Szilagyi, L., Rutter, W.J., 1992. Converting trypsin to chymotrypsin: the role of surface loops. Science 255, 1249–1253.

Hilder, V.A., Boulter, D., 1999. Genetic engineering of crop plants for insect resistance—a critical review. Crop Protec. 18, 177–191.

Igaz, P., Fejér, G., Szalai, C., Tóth, S., Falus, A., 1998. Development of competitive mRNA PCR for the quantification of interleukin-6-responsive *junB* oncogene expression. BioTechniques 24, 854–859.

Jongsma, M.A., Bakker, P.L., Peters, J., Bosch, D., Stiekema, W.J., 1995. Adaptation of *Spodoptera exigua* larvae to plant proteinase inhibitors by induction of gut proteinase activity insensitive to inhibition. Proc. Natl. Acad. Sci. 92, 8041–8045.

Kraut, J., 1977. Serine proteases: structure and mechanism of catalysis. Ann. Rev. Biochem. 46, 331–358.

Martínez-Ramírez, A.C., Real, M.D., 1996. Proteolytic processing of *Bacillus thuringiensis* CryIIIA toxin and specific binding to brush-border membrane vesicles of *Leptinotarsa decemlineata* (Colorado potato beetle). Pestic. Biochem. Physiol. 54, 115–122.

McGaughey, W.H., 1985. Insect resistance to the biological insecticide *Bacillus thuringiensis*. Science 229, 193–195.

McGaughey, W.H., Beeman, R.W., 1988. Resistance to *Bacillus thuringiensis* in colonies of Indianmeal moth and almond moth (Lepidoptera: Pyralidae). J. Econ. Entomol. 81, 28–33.

McGaughey, W.H., Johnson, D.E., 1992. Indianmeal moth (Lepidoptera: Pyralidae) resistance to different strains and mixtures of *Bacillus thuringiensis*. J. Econ. Entomol. 85, 1594–1600.

Michaud, D., Cantin, L., Vrain, T.C., 1995. Carboxyl-terminal truncation of oryzacystatin II by oryzacystatin-insensitive insect digestive proteinases. Arch. Biochem. Biophys. 322, 469–474.

Milne, R., Kaplan, H., 1993. Purification and characterization of a trypsin-like digestive enzyme from spruce budworm (*Choristoneura fumiferana*) responsible for the activation of δ-endotoxin from *Bacillus thuringiensis*. Insect Biochem. Molec. Biol. 23, 663–673.

Müller, H.M., Crampton, J.M., Del La Torre, A., Sinden, R., Crisanti, A., 1993. Members of a trypsin gene family in *Anopheles gambiae* are induced in the gut by blood meal. EMBO J. 12, 2891–2900.

Oppert, B., Kramer, K.J., Johnson, D.E., MacIntosh, S.C., McGaughey, W.H., 1994. Altered protoxin activation by midgut enzymes from a *Bacillus thuringiensis* resistant strain of *Plodia interpunctella*. Biochem. Biophys. Res. Commun. 198, 940–947.

Oppert, B., Kramer, K.J., Johnson, D.E., Upton, S.J., McGaughey, W.H., 1996. Luminal proteinases from *Plodia interpunctella* and the hydrolysis of *Bacillus thuringiensis* CryIA(c) protoxin. Insect Biochem. Molec. Biol. 26, 571–583.

Oppert, B., Kramer, K.J., Beeman, R.W., Johnson, D., McGaughey, W.H., 1997. Proteinase-mediated insect resistance to *Bacillus thuringiensis* toxins. J. Biol. Chem. 272, 23473–23476.

Pereira, P.J., Bergner, A., Macedo-Ribeiro, S., Huber, R., Matschiner, G., Fritz, H., Sommerhoff, C.P., Bode, W., 1998. Human beta-tryptase is a ring-like tetramer with active sites facing a central pore. Nature 392, 306–311.

Peterson, A.M., Barillas-Mury, C.V., Wells, M.A., 1994. Sequence of three cDNAs encoding an alkaline midgut trypsin from *Manduca sexta*. Insect Biochem. Molec. Biol. 24, 463–471.

Peterson, A.M., Fernando, G.J.P., Wells, M.A., 1995. Purification, characterization and cDNA sequence of an alkaline chymotrypsin from the midgut of *Manduca sexta*. Insect Biochem. Molec. Biol. 25, 765–774.

Reeck, G., Oppert, B., Denton, M., Kanost, M., Baker, J., Kramer, K., 1999. Insect proteinases. In: Turk, V. (Ed.), Proteases—New Perspectives, Birkhäuser, Basel, Switzerland, (in press).

Sasaki, T., Hishida, T., Ichikawa, K., Asari, S., 1993. Amino acid sequence of alkaliphilic serine protease from silkworm *Bombyx mori*, larval digestive juice. FEBS Lett. 320, 35–37.

Shao, Z., Cui, Y., Liu, X., Yi, H., Ji, J., Yu, Z., 1998. Processing of delta-endotoxin of *Bacillus thuringiensis* subsp. *kurstaki* HD-1 in *Heliothis armigera* midgut juice and the effects of protease inhibitors. J. Invert. Pathol. 72, 73–81.

Terra, W.R., Ferriera, C., 1994. Insect digestive enzymes: properties, compartmentalization and function. Comp. Biochem. Physiol. 109B, 1–62.

Titus, D., 1991. Promega Protocols and Application Guide, 2nd ed. Promega Corporation, Madison, WI.

Valaitis, A.P., Augustin, S., Clancy, K.M., 1999. Purification and characterization of the western spruce budworm larval midgut proteinases and comparison of gut activities of laboratory-reared and field-collected insects. Insect Biochem. Molec. Biol. 29, 405–414.

Wang, S., Magoulas, C., Hickey, D.A., 1993. Isolation and characterization of a full-length trypsin-encoding cDNA clone from the lepidopteran insect *Choristoneura fumiferana*. Gene 136, 375–376.

Wang, S., Young, F., Hickey, D.A., 1995. Genomic organization and expression of a trypsin gene from the spruce budworm *Choristoneura fumiferana*. Insect Biochem. Molec. Biol. 25, 899–908.

Wang, S., Magoulas, C., Hickey, D., 1999. Concerted evolution within a trypsin gene cluster in *Drosophila*. Molec. Biol. Evol. 16, 1117–1124.

Zhu, Y.C., Oppert, B., Kramer, K.J., McGaughey, W.H., Dowdy, A.K., 1997. cDNAs for a chymotrypsinogen-like protein from two strains of *Plodia interpunctella*. Insect Biochem. Molec. Biol. 27, 1027–1037.

Zhu, Y.C., Oppert, B., Kramer, K.J., McGaughey, W.H., Dowdy, A.K., 2000. cDNA sequences, mRNA expression levels, and genomic DNAs of trypsinogen from the Indianmeal moth *Plodia interpunctella*. Insect Molec. Biol. 9, 19–26.